

# Global disorder transition in the community structure of large- $q$ Potts systems

PETER RONHOVDE, DANDAN HU and ZOHAR NUSSINOV<sup>(a)</sup>

*Department of Physics, Washington University in St. Louis - Campus Box 1105, 1 Brookings Drive, St. Louis, MO 63130, USA*

PACS 89.75.Fb – Structures and organization in complex systems  
PACS 64.60.Cn – Order-disorder transformations  
PACS 64.60.De – Statistical mechanics of model systems

**Abstract** – We examine a global disorder transition when identifying community structure in an arbitrary complex network. Earlier, we illustrated [*Phil. Mag.* **92**, 406 (2012)] that “community detection” (CD) generally exhibits disordered (or unsolvable) and ordered (solvable) phases of both high and low computational complexity along with corresponding transitions from regular to chaotic dynamics in derived systems. Using an exact generalized dimensional reduction inequality, multivariate Tutte polynomials, and other considerations, we illustrate how increasing the number of communities  $q$  emulates increasing the heat bath temperature  $T$  for a *general* weighted Potts model, leading to global disorder in the community structure of *arbitrary* large graphs. Dimensional reduction bounds lead to results similar to those suggested by mean-field type approaches. Large systems tend toward global insolubility in the limit of large  $q$  above a crossover temperature  $T_{\times} \approx L|J_e|/[N \ln q]$  where  $|J_e|$  is a typical interaction strength,  $L$  is the number of edges, and  $N$  is the number of nodes. For practical system sizes, a solvable phase is generally accessible at low  $T$ . The global nature of the disorder transition does not preclude solutions by local CD algorithms (even those that employ global cost function parameters) as long as community evaluations are locally determined.

**Introduction.** – Methods of statistical physics have enlarged the understanding of complex networks [1]. In particular, community detection (CD) [2] attempts to identify “mesoscopic” structure within these systems. Applications of CD are extremely broad, and numerous methods have been leveraged to solve it [3–14]. The problem complexity and related aspects of community “detectability” were studied for an “absolute Potts model” (APM) [15, 16], modularity [17, 18], and mean-field type (cavity) approaches [19, 20] where the latter references [18–20] examined general cluster detectability transitions in a special class of stochastic block models. Ref. [21] reviewed critical phenomena in complex networks.

We illustrated [16] that the APM, along with other CD approaches, exhibits solvable phases of “easy” or “hard” complexity and unsolvable phases with spin-glass-type or other transitions coinciding with transitions from ergodic to non-ergodic dynamics in mechanical analogs. Bashan *et al.* [22] showed that transitions in network topology

have physiological significance, implying that network transitions have relevance beyond detectability/solvability thresholds. We also previously demonstrated [23] how distinct phases of the CD problem affect image segmentation applications. Other authors covered disorder transitions for random-bond Potts models [24, 25] and zeros of the partition function [26] in the limit of a large number of Potts spin flavors ( $q \gg 1$ ).

As depicted in fig. 1, CD attempts to partition a graph into  $q$  optimally disjoint subgraphs (or communities). Optimal values of  $q$  may be determined via multiscale methods [9, 10, 27–29]. Ref. [30] discussed the absence of large clusters in large real networks (*i.e.*,  $q$  is large).

We investigate a *general* weighted Potts model on an *arbitrary* graph with  $q \gg 1$ , and we illustrate how increasing  $q$  emulates increasing the temperature  $T$ . For CD, it implies that whenever an algorithm or cost function can be represented as a weighted Potts model, then *large systems are inherently disordered* on a global level above a crossover temperature  $T_{\times}$ . The result encom-

<sup>(a)</sup>E-mail: zohar@wuphys.wustl.edu

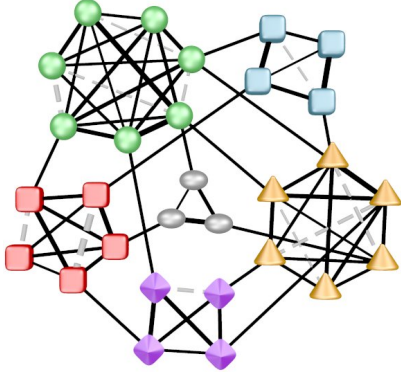


Fig. 1: (Color online) The schematic illustrates a community partition (distinct node shapes and colors) showing relevant structure in a graph with ferromagnetic (solid, black lines) and antiferromagnetic interactions (gray, dashed lines). Line thickness indicates the relative interaction strength  $J_e$ . Antiferromagnetic (“adversarial” with  $J_e < 0$ ) and non-interacting (“neutral,” unconnected with  $J_e = 0$ ) relations break up well-defined communities.

passes a wide variety of CD methods including optimizing modularity [31], a Potts model applying a “configuration null model” (CMPM) [4, 32], an Erdős-Rényi Potts model [4, 33], a “constant Potts model” [34], “label propagation” [11, 35], the APM [10, 15], and others [36–38]. We speculate that the disorder persists for systems with external magnetic interactions [39] as well as directed and multipartite graphs.

While the result applies to general Potts models on arbitrary large graphs, it only implies disorder on a global, as opposed to local, level for large- $q$  networks with bounded coordination numbers and vertex count. Refs. [15, 34] were shown to avoid a “resolution limit” imposed by global *cost function* parameters on some models [4, 31, 40–42], but all weighted (or unweighted) Potts models would be subject to the *global* disorder imposed by large  $q$ . This global disorder can be mitigated or avoided by solving the system locally or at sufficiently low  $T$ , but the latter condition exists in competition with beneficial thermal annealing effects of increased temperature at sufficiently low  $T$  (“order by disorder”) [16, 43].

The disorder induced by large  $q$  is quantitatively different from that caused by high noise (extraneous inter-community edges) in a network [15, 16, 43]. A “glassy” transition due to noise may persist as  $T \rightarrow 0$  because the solution algorithm is frustrated by a complex energy landscape exhibiting numerous local minima.

**Potts Hamiltonian.** — We consider a general spin-glass-type Potts model Hamiltonian

$$H(\{\sigma\}) = - \sum_{i \neq j} J_{ij} \delta(\sigma_i, \sigma_j), \quad (1)$$

where  $J_{ij}$  is the interaction strength between spins  $i$  and  $j$  and  $\delta(\sigma_i, \sigma_j) = 1$  if  $\sigma_i = \sigma_j$  and 0 otherwise. For CD,

it is convenient separate the ferromagnetic ( $J_{ij} > 0$ ) and antiferromagnetic ( $J_{ij} \leq 0$ ) interactions

$$H(\{\sigma\}) = - \sum_{i \neq j} [w_{ij} A_{ij} - u_{ij} (1 - A_{ij})] \delta(\sigma_i, \sigma_j). \quad (2)$$

Given  $N$  nodes,  $\{A_{ij}\}$  is the adjacency matrix where  $A_{ij} = 1$  if nodes  $i$  and  $j$  are connected by a ferromagnetic edge and is 0 otherwise.  $w_{ij} > 0$  and  $u_{ij} \geq 0$  are ferromagnetic and antiferromagnetic edge weights, respectively. Each spin  $\sigma_i$  may assume integer values in the range  $1 \leq \sigma_i \leq q$  where  $q$  is usually dynamically determined. Node  $i$  is a member of community  $k$  when  $\sigma_i = k$ .

The antiferromagnetic weights  $u_{ij}$  provide a “penalty function” enabling a non-trivial CD ground state for an arbitrary graph. Some models incorporate a weight factor, generally on the  $u_{ij}$  term, which allows the model to span different network scales in qualitatively similar ways. The APM penalizes “neutral” relationships (*i.e.*, generally  $u_{ij} \equiv J_{ij} - 1$  for  $J_{ij} \leq 0$ ). Another model [34] incorporates weighted antiferromagnetic interactions into  $w_{ij}$  and applies a separate penalty term. Algorithms for modularity and the CMPM are effectively implemented with dynamic edge weights on the  $u_{ij}$  term, but the fluctuations would be small in general approaching the ground state. Local CD models for energy calculations were suggested in [40, 41], further advocated in [15], and explored in more detail in [34].

**Dimensional reduction bound.** — We first provide rigorous bounds on the disorder transition for community structure using dimensional reduction inequalities [44, 45]. In the current context, these simple, yet exact, inequalities relate a system in any dimension to a local ( $D = 0$  dimensional) system composed of a single vertex (or a finite collection of vertices) and its (their) neighbors. The derived bound has a form similar to that suggested by mean field considerations.

In the thermodynamic limit, a bona fide transition may occur that marks symmetry breaking wherein, for infinitesimally weak applied fields that favor a particular state, the probability that a given spin belongs to one of the  $q$  communities differs from  $1/q$ . From a practical standpoint, we are interested in the probability that a particular spin  $\sigma_0$  takes on a specific “correct” spin value  $\bar{\sigma}$  that it does in a low energy configuration, effectively searching for a “planted” solution  $\bar{\sigma}$ .

We derive upper bounds on the temperature for which the spin  $\sigma_0$  attains its correct value with high confidence. Towards this end, we first detail general inequalities and then turn to their application in our case. We consider a partition of all spins into those of a local set  $\eta$  (*i.e.*,  $\sigma_0$  in the single-spin case) and all other remaining spins  $\psi$  in the system. The trace over all spins becomes  $\text{Tr}_{\{\sigma\}} = \text{Tr}_{\{\psi\}} \text{Tr}_{\{\eta\}}$  and the Hamiltonian (with or without any weak applied fields) becomes a function involving both sets of spins  $H(\{\sigma\}) = H(\{\psi\}, \{\eta\})$ . Any thermal

average  $\langle f(\eta) \rangle$  can be written as

$$\langle f \rangle = \frac{\text{Tr}_{\{\psi\}} z_\psi \left[ \frac{\text{Tr}_{\{\eta\}} f(\eta) e^{-\beta H(\psi, \eta)}}{z_{\{\psi\}}} \right]}{\text{Tr}_{\{\psi\}} z_\psi}. \quad (3)$$

where we inserted  $z_\psi \equiv \text{Tr}_{\{\eta\}} e^{-\beta H(\psi, \eta)} \geq 0$  twice in the numerator which is valid for any Hamiltonian. As can be readily seen, the exact  $\langle f \rangle$  over the large system can be written as a weighted sum (with positive normalized weights,  $p_\psi = z_\psi / \text{Tr}_{\{\psi'\}} z_{\psi'}$ , that sum to unity) of local averages

$$\langle f \rangle_\psi \equiv \frac{\text{Tr}_{\{\eta\}} f(\eta) e^{-\beta H(\psi, \eta)}}{z_\psi}. \quad (4)$$

Thus,  $\langle f \rangle \leq \langle f \rangle_{\bar{\psi}}$ , where  $\bar{\psi}$  is a particular set of the spins  $\psi$  that maximizes  $\langle f \rangle_\psi$ . When we substitute  $\psi = \bar{\psi}$  in  $H$ , we obtain a local Hamiltonian  $H(\bar{\psi}, \eta)$  in the spins  $\eta$ .

If we set  $f \equiv \delta(\sigma_0, \bar{\sigma})$ , then the mean value of  $\langle f \rangle$  will correspond to the probability  $\langle f \rangle = P(\sigma_0 = \bar{\sigma})$ . When computing the internal general trace over  $\eta$ , we evaluate  $\langle f \rangle_{\bar{\psi}}$ , in the case of a single spin at the origin  $\sigma_0$  averaged over its  $q$  possible states. Applying  $\langle f \rangle \leq \langle f \rangle_{\bar{\psi}}$  to  $\langle \delta(\sigma_0, \bar{\sigma}) \rangle$ , we obtain a generous upper bound. If the interaction between  $\sigma_j$  at site  $j$  and  $\sigma_0$  at the origin is larger than 0 (*i.e.*,  $J_{j0} > 0$ ), then  $\bar{\psi}_j = \bar{\sigma}$ . If  $J_{j0} < 0$ , then we set  $\bar{\psi}_j \neq \bar{\sigma}$ . With this set of  $\psi$  values,

$$p = \langle f \rangle \leq \frac{e^{\beta \bar{J}_0}}{e^{\beta \bar{J}_0} + (q-1)}, \quad (5)$$

where  $\bar{J}_0 = \frac{1}{2} \sum_j J_{j0} [1 + \text{sgn}(J_{j0})]$ .

On practical benchmarks, we are interested in cases where  $p$  exceeds some threshold value  $p^*$ . The inverse temperature  $\beta^*$  at which the probability exceeds  $p^*$  is

$$\beta^* = \frac{1}{\bar{J}_0} \ln \left[ \frac{p^*(q-1)}{(1-p^*)} \right]. \quad (6)$$

At high  $q$  with  $p^* = 1/2$ , this leads to rigorous upper bound (UB) for the associated crossover temperature

$$T_\times^{\text{UB}} = \frac{\bar{J}_0}{k_B \ln q}. \quad (7)$$

For  $T > T_\times^{\text{UB}}$ , the correct assignment can only be determined with a probability  $p^* = 1/2$ . If the exchange constants  $J_{j0}$  and the coordination number of  $\sigma_0$  are finite and do not match or exceed the  $\ln q$  dependence in the denominator, the system is unsolvable at any positive temperature as  $q \rightarrow \infty$ . The bound of eq. (7) for the node at the origin is local, so it may change for other nodes.

In practice, some parts of the network can exhibit structure at higher temperatures which serves as a bottleneck for global ordering. Generally, the bounds of  $\langle f \rangle \leq \langle f \rangle_{\bar{\psi}}$  enable a reduction of the full physical system to a related problem that occupies a reduced  $D$ -dimensional sub-volume of the entire system. If we define the external state

$\psi$  as a set of spins with the average spin value, then the resulting average becomes a mean-field average.

We discuss a general representation for Potts model where, with it and related approaches, we estimate the form of the cross-over (or transition) temperature  $T_\times$  from a viable low temperature ordered phase to a high temperature disordered regime. The scaling in these results is similar to that of the rigorous bound in eq. (7).

**Multivariate Tutte polynomial estimate.** – The multivariate Tutte polynomial [46] is defined as a sub-graph expansion over  $\mathcal{A} \subseteq \mathcal{E}$  of a graph  $G = (V, \mathcal{E})$  where  $V$  and  $\mathcal{E}$  are the sets of vertices and (ferromagnetic and antiferromagnetic) edges, respectively.

$$Z(G; q, \mathbf{v}) = \sum_{\mathcal{A} \subseteq \mathcal{E}} q^{k(\mathcal{A})} \prod_{e' \in \mathcal{A}} v_{e'} \quad (8)$$

where  $k(\mathcal{A})$  is the number of connected components of  $G_{\mathcal{A}} = (V, \mathcal{A})$ ,  $v_e = \exp(\beta J_e) - 1$ , and  $J_e$  is the interaction strength of edge  $e$ . In CD, large  $q$  necessarily implies a large number of nodes  $|V| = N$ .

For two disjoint partitions  $A$  and  $B$  with  $G = A \cup B$ ,  $Z(G; q, \mathbf{v}) = Z(A; q, \mathbf{v}_A) Z(B; q, \mathbf{v}_B)$  where  $\mathbf{v}_A$  and  $\mathbf{v}_B$  are the edge weights in the respective subgraphs. For unweighted systems, the interaction strength is  $J_e \equiv J = \pm 1$  where  $+$  and  $-$  correspond to ferromagnetic or antiferromagnetic interactions, respectively.

There is a slight terminology distinction between CD and the energy contributions in eq. (1). Edges with  $J_e > 0$  correspond to the  $w_{ij}$  ferromagnetic (“friendly” or “cooperative”) interactions in eq. (2), and  $J_e \leq 0$  relates to the  $u_{ij}$  anti-ferromagnetic (neutral and “adversarial” in some models) or absent (neutral in some models) interactions. The edge effect is conceptually consistent with CD for  $J_e > 0$ , but antiferromagnetic weights are also related by an edge when calculating eq. (8). That is, an interaction exists, but it is antiferromagnetic in nature. In CD, repulsive antiferromagnetic interactions correspond to adversarial relationships which act like neutral (unconnected) relations that hinder community structure.

For large  $T$ , we require  $T \gg \max_{e \in \mathcal{E}} |J_e|$ . The leading order terms for an *arbitrary* graph are due to  $\mathcal{A}_\emptyset = \{\emptyset\}$  and  $\mathcal{A}_e = \{e\}$  for each edge  $e \in \mathcal{E}$ . We also include the last  $\mathcal{A} = \mathcal{E}$  term of  $G$  which is addressed later in the text.

$$Z(G; q, \mathbf{v}) = q^N \left( 1 + \sum_{e'=1}^{|\mathcal{E}|} \frac{v_{e'}}{q} + \cdots + q^{k(G)-N} \prod_{f'=1}^{|\mathcal{E}|} v_{f'} \right) \quad (9)$$

which applies to arbitrarily large systems of size  $N$ . The free energy per site  $f = -\frac{k_B T}{N} \ln Z$  becomes

$$f \approx -k_B T \ln q - \frac{k_B T}{N} \sum_{e'=1}^{|\mathcal{E}|} \frac{v_{e'}}{q} \quad (10)$$

where we invoke the small  $x$  approximation,  $\ln(1+x) \approx x$ , and neglect the last  $\mathcal{A} = \mathcal{E}$  term in eq. (9). For large  $T$ ,

$v_e \approx J_e/(k_B T)$  and

$$f \approx -k_B T \ln q - \frac{1}{N} \sum_{e'=1}^{|\mathcal{E}|} \frac{J_{e'}}{q}. \quad (11)$$

When compared to eq. (15) for large  $q$ , it implies that  $q$  emulates  $T$  when  $T$  is large.

When  $q \gg \max_{e \in \mathcal{E}} |v_e|$ , the leading order terms in  $1/q$  are still the first two in eq. (9). As  $T \rightarrow 0$ , higher order terms in  $v_e$  become increasingly important which is dominated by the last subgraph term for  $\mathcal{A} = \mathcal{E}$ . The three displayed terms are “universal” since they apply equally well to all graphs (*e.g.*, lattices, Cayley trees, random graphs, etc.). That is, they only depend on the system size  $N$ , the number of links  $L$ , and the number of connected components  $k(G)$  for the full graph  $G$ .

All displayed terms in eq. (9) are identical for regular lattices and similar-coordination-number Bethe lattices. Similar results are obtained for other graphs where Bethe lattice approximants are only identical for these terms. The important non-universal terms in the subgraph expansion [denoted by ellipsis] depend on the particular graph structure. For the leading and last terms, respectively, the logarithms of the terms flesh out the typical scales of the high temperature entropic and low temperature energetic contributions to the free energy.

In the large- $q$  limit, the zeroes of  $Z$  for constant  $v$  provide the relevant transition temperatures in the  $N \rightarrow \infty$  limit, and the free energy per site becomes non-analytic. We can estimate disorder transition temperatures by comparing the second and last terms,  $q^{N-1} \sum_{e'} v_{e'}$  to  $q^{k(G)} \prod_{f'=1} v_{f'}$ , assuming a “typical” interaction strength  $|J_e|$  so that we can solve the equation. If  $q$  is large then the latter term in eq. (9) will compete with the last term suggesting a crossover temperature

$$T_\times \approx \frac{|J_e|}{k_B \ln (q^{[N-k(G)]/L} + 1)} \quad (12)$$

under the assumptions  $N \gg 1$  and  $L \gg 1$ . For general  $\{J_e\}$ , we may see multiple transitions spread over a range of  $T$ . In the limit as  $T \rightarrow 0$ , eq. (12) becomes

$$T_\times \approx \frac{L|J_e|}{k_B [N - k(G)] \ln q}. \quad (13)$$

If we instead compare  $v_{f'}/q$  to 1 in eq. (9), the factor  $L/N$  disappears, but the logarithmic behavior in  $q$  remains. Eq. (13) diverges for an arbitrarily large complete graph [ $L = N(N-1)/2$  and  $N \rightarrow \infty$ ], and it approaches zero as  $q \rightarrow \infty$ . Often in CD, the graph is (almost) completely connected (in a topological sense) so  $N \gg k(G)$ . For *sparse* graphs,  $L \propto N$ , so

$$T_\times^{\text{Sparse}} \approx \frac{d|J_e|}{2k_B \ln q} \quad (14)$$

where  $d$  is an average node degree.

Above  $T_\times$ , the large- $q$  contributions dominate, and the system is in a disordered state, but it is globally ordered for  $T \ll T_\times$ . For moderate levels of noise in a graph, larger  $d$  (with a well-defined community structure) actually *increases* the crossover temperature. This indicates that additional noise up to an insurmountable threshold allows the system to explore the phase space more completely when the community structure is solved [43]. For degree distributions seen in CD (*e.g.*, often a power law), the corresponding crossover temperature would spread or split into multiple values which model the distinct features of the graph. In the limit of large  $N$ , the crossover(s) would become an approximate transition point.

In order to highlight the similarity between the large- $q$  and  $T$  behaviors, we fix  $T = T' > T_\times$ , define the constant  $J_e^{(q)} \equiv k_B T' \{\exp [J_e/(k_B T')]\} - 1$ , and rewrite eq. (10) as

$$f \approx -k_B T' \ln q - \frac{1}{N} \sum_{e'=1}^{|\mathcal{E}|} \frac{J_{e'}^{(q)}}{q}. \quad (15)$$

Large  $q$  in eq. (15) emulates large  $T$  in eq. (11).  $J_e^{(q)}$  is exponentially weighted in  $\beta'$ , so a non-zero (perhaps small) region of stability is ensured except in the presence of high noise [15, 16, 43]. Transitions between contending minima in random embedded systems [16], contending states with multi-scale (*e.g.*, hierarchical) structures [10], or others may occur over a range of temperatures.

**Mean-field and free energy estimates.** – The mean-field transition temperature for lattices with a fixed coordination number  $d$ , constant exchange  $J$ , and arbitrary  $q$  is [47, 48]

$$T_c^{\text{MF}} = \frac{Jd(q-2)}{2k_B(q-1) \ln(q-1)} \quad (16)$$

for  $q \geq 3$ . This equation yields a large- $q$  limit of

$$T_c^{\text{MF}} \approx \frac{dJ}{2k_B \ln q} \quad (17)$$

in agreement with eq. (14). The  $q \rightarrow \infty$  limit on ferromagnetic lattices asymptotically approaches the mean-field theory result with translationally invariant  $J$  and constant  $d$  [47, 49]. The Gibbs-Bogoliubov-Feynman inequality also allows a method for deriving optimal mean-field approximations in general.

We can ascertain the same asymptotic behavior in  $q$  by analyzing the free energy per site if we flip a spin in a ground state. Assuming that the energy and entropy changes are uncorrelated, the energy change for the node flip is  $\Delta U \simeq d|J_e|$  up to an undetermined constant factor, and the entropy change is  $\Delta S \simeq \ln q$  yielding a free energy change  $\Delta F \simeq d|J_e| - k_B T \ln q$ . The entropy contribution dominates (see [26] on general lattices)  $\Delta F$  above a crossover estimate

$$T_\times^{\text{FE}} \approx \frac{d|J_e|}{k_B \ln q} \quad (18)$$

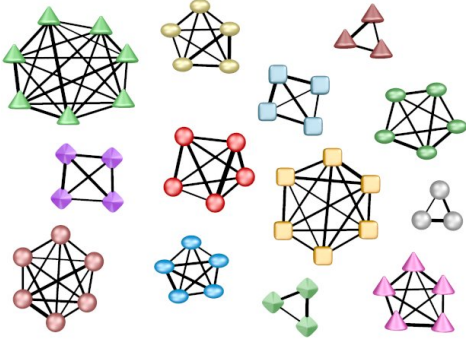


Fig. 2: (Color online) The figure depicts  $q$  independent maximal sub-graphs of size  $n_i$  for  $i = 1$  to  $q$ . The line thickness illustrates the relative strengths. With strictly ferromagnetic weights as depicted, this system is the *strongest* possible community structure since it is “perfectly” defined via a maximal number of internal edges with no external edges to muddle the natural structure. Any antiferromagnetic weights would weaken the community structure much like “neutral” relations act to break up a well-defined community. In eqs. (7)–(18), we show that *all* large- $q$  CD problems that are representable as a general weighted Potts model experience *global* disorder at high  $q$  above a crossover temperature  $T_\times$ , and we calculate the partition function and the free energy per site of this system as a specific example in eqs. (19)–(21).

which agrees well with eq. (14). As we alluded earlier, the logarithms of the leading and last terms in eq. (9) trivially provide the typical entropic and energetic contributions to the free energy at high and low temperatures.

**Non-interacting cliques example.** – The *most* strongly defined community structure is a system of  $q$  non-interacting cliques (maximally connected sub-graphs with no intercommunity relationships that obscure community structure). We define  $q$  weighted cliques with sizes  $n_i$  for  $i = 1$  to  $q$  as depicted in fig. 2. The partition function at high  $T$  or high  $q$  with  $T \gg T_\times$  is [43]

$$Z(G; q, \mathbf{v}) \approx q^N \prod_{i=1}^q \prod_{j=1}^{n_i} \left( 1 + \sum_{k=1}^{j-1} \frac{v_{k+\ell_j}}{q} \right) \quad (19)$$

where  $\ell_j = (j-1)(j-2)/2$ . Eq. (19) is equivalent to eq. (9) to first order in  $v_k$ . When  $T \gg T_\times$ , high  $q$  results in the same approximation as high  $T$  affirming the implication made with eqs. (11) and (15).

For high  $T$  specifically, we make the additional approximation  $v_k \approx \beta J_k$ , and the free energy per site becomes

$$f \approx -k_B T \ln q - \frac{E_N}{q} \quad (20)$$

using the same approximations as in eq. (10).  $E_N = \sum_{e'} J_{e'}/N$  is the energy per site. In ref. [43], we derived

$$T_\times^{\text{NIC}} \approx \frac{(n-1)J}{2k_B \ln q} \quad (21)$$

for constant  $J$  and  $q$  non-interacting cliques of *fixed* size  $n$ . With  $d = n - 1$  edges per node, the result coincides well with eqs. (14), (17) and (18).

**Thermal annealing comments.** – For heat bath or simulated annealing (SA) algorithms in CD, when  $T \ll L|J_e|/[Nk_B \ln q]$ , the global system remains in an ordered state. For higher  $T$ , the *global* system becomes increasingly disordered in terms of partition function state probabilities. If we consider states “near” equilibrium, small fluctuations in the energy result in only tiny changes to the state probabilities through the Boltzmann weight. Another perspective is that even for a system of non-interacting cliques depicted in fig. 2, larger  $q$  creates a greater probability that a non-negligible fraction of cliques will be disconnected at a given  $T$  at any given point during the stochastic solution.

Most SA algorithms (*e.g.*, [4]) utilize energy *differences* to evaluate dynamic changes to the community division, so the system is effectively solved locally (algorithmically speaking, global parameters in the *cost function* are a separate issue [4, 15, 31, 34, 40–42, 50]). In practice, SA is limited to systems with  $O(10^4)$  nodes without significant parallelization, so greedy algorithms ( $T = 0$ ) are used on the largest systems [5, 15, 35]. SA implements a cooling scheme to fine tune solutions determined by the high  $T$  optimization, and our results indicate that cooling becomes more important for large- $q$  systems.

**Conclusion.** – We showed a global disorder transition at a large number of communities  $q$  for a *general* weighted (or unweighted) Potts model over essentially *arbitrary* graphs. The community structure of a complex network may be *globally* disordered at large  $q$  but still be *locally* ordered and locally solvable. Our results encompass many popular cost functions utilized for community detection, including modularity and common Potts model variants. We demonstrated this effect using stringent exact bounds as well as related results suggested by mean-field and other general approaches. With these bounds, results for a local system that occupies only a sub-volume of the original system lead to rigorous results for the full system, and they may have similar applications in the analysis of other hard computational problems where mean-field approaches are commonly applied. We also illustrated that in the strongest possible model partition, that of non-interacting cliques, the large- $q$  limit induces disorder akin to random thermal effects.

Increasing  $q$  emulates increasing  $T$  in arbitrary graphs for any CD method that may be represented as a general weighted Potts model. The asymptotic behavior of the global disorder transition varies slowly in  $q$ ,  $T_\times \approx L|J_e|/[Nk_B \ln q]$ , meaning that problems of practical size maintain a finite region of solvability given a stochastic heat bath algorithm. Local algorithm dynamics (even for models which incorporate global weighting parameters) serve to circumvent the global disorder transition. This global disorder is generally circumvented by the of-

ten used SA algorithm, but “glassy” problems with high noise (many extraneous intercommunity edges) would remain a challenge for any algorithm or model.

\* \* \*

This work was supported by NSF grant DMR-1106293 (ZN). We also wish to thank M. Biskup, S. Chakrabarty, L. Chayes, V. Dobrosavljevic, P. Johnson, and L. Zdeborová for discussions and ongoing work.

## REFERENCES

- [1] NEWMAN M. E. J., *Phys. Today*, **61** (2008) 33.
- [2] FORTUNATO S., *Phys. Rep.*, **486** (2010) 75.
- [3] CLAUSET A., NEWMAN M. E. J. and MOORE C., *Phys. Rev. E*, **70** (2004) 066111.
- [4] REICHARDT J. and BORNHOLDT S., *Phys. Rev. E*, **74** (2006) 016110.
- [5] BLONDEL V. D., GUILLAUME J.-L., LAMBIOTTE R. and LEFEBVRE E., *J. Stat. Mech.*, **10** (2008) P10008.
- [6] ROSVALL M. and BERGSTROM C. T., *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008) 1118.
- [7] GUDKOV V., MONTEALEGRE V., NUSSINOV S. and NUSSINOV Z., *Phys. Rev. E*, **78** (2008) 016113.
- [8] KUMPULA J. M., KIVELÄ M., KASKI K. and SARAMÄKI J., *Phys. Rev. E*, **78** (2008) 026109.
- [9] LANCICHINETTI A., FORTUNATO S. and KERTÉSZ J., *New J. Phys.*, **11** (2009) 033015.
- [10] RONHOVDE P. and NUSSINOV Z., *Phys. Rev. E*, **80** (2009) 016109.
- [11] BARBER M. J. and CLARK J. W., *Phys. Rev. E*, **80** (2009) 026129.
- [12] LANCICHINETTI A. and FORTUNATO S., *Phys. Rev. E*, **80** (2009) 056117.
- [13] CHENG X.-Q. and SHEN H.-W., *J. Stat. Mech.*, **04** (2010) P04024.
- [14] SHEN H.-W. and CHENG X.-Q., *J. Stat. Mech.*, **10** (2010) P10020.
- [15] RONHOVDE P. and NUSSINOV Z., *Phys. Rev. E*, **81** (2010) 046114.
- [16] HU D., RONHOVDE P. and NUSSINOV Z., *Phil. Mag.*, **92** (2012) 406.
- [17] GOOD B. H., DE MONTJOYE Y.-A. and CLAUSET A., *Phys. Rev. E*, **81** (2010) 046106.
- [18] NADAKUDITI R. R. and NEWMAN M. E. J., *Phys. Rev. Lett.*, **108** (2012) 188701.
- [19] REICHARDT J. and LEONE M., *Phys. Rev. Lett.*, **101** (2008) 078701.
- [20] DECELLE A., KRZAKALA F., MOORE C. and ZDEBOROVÁ L., *Phys. Rev. Lett.*, **107** (2011) 065701.
- [21] DOROGVTSEV S. N., GOLTSEV A. V. and MENDES J. F. F., *Rev. Mod. Phys.*, **80** (2008) 1275.
- [22] BASHAN A., BARTSCH R. P., KANTELHARDT J. W., HAVLIN S. and IVANOV P. C., *Nature Comm.*, **3** (2012) 702.
- [23] HU D., RONHOVDE P. and NUSSINOV Z., *Phys. Rev. E*, **85** (2012) 016101.
- [24] JUHÁSZ R., RIEGER H. and IGLÓI F., *Phys. Rev. E*, **64** (2001) 056122.
- [25] MERCALDO M. T., D’AURIAC J.-C. A. and IGLÓI F., *Europhys. Lett.*, **70** (2005) 733.
- [26] CHANG S.-C. and SHROCK R., *Int. J. Mod. Phys. B*, **21** (2007) 979.
- [27] ARENAS A., FERNÁNDEZ A. and GÓMEZ S., *New J. Phys.*, **10** (2008) 053039.
- [28] KUMPULA J. M., SARAMÄKI J., KASKI K. and KERTÉSZ J., *Fluct. Noise Lett.*, **7** (2007) L209.
- [29] ROSVALL M. and BERGSTROM C. T., *PLoS ONE*, **6** (2011) e18209.
- [30] LESKOVEC J., LANG K. J., DASGUPTA A. and MAHONEY M. W., *Internet Mathematics*, **6** (2009) 29.
- [31] NEWMAN M. E. J. and GIRVAN M., *Phys. Rev. E*, **69** (2004) 026113.
- [32] TRAAG V. A. and BRUGGEMAN J., *Phys. Rev. E*, **80** (2009) 036115.
- [33] REICHARDT J. and BORNHOLDT S., *Phys. Rev. Lett.*, **93** (2004) 218701.
- [34] TRAAG V. A., VAN DOOREN P. and NESTEROV Y., *Phys. Rev. E*, **84** (2011) 016114.
- [35] RAGHAVAN U. N., ALBERT R. and KUMARA S., *Phys. Rev. E*, **76** (2007) 036106.
- [36] BLATT M., WISEMAN S. and DOMANY E., *Phys. Rev. Lett.*, **76** (1996) 3251.
- [37] ISPOLATOV I., MAZO I. and YURYEV A., *J. Stat. Mech.*, **09** (2006) P09014.
- [38] HASTINGS M. B., *Phys. Rev. E*, **74** (2006) 035102.
- [39] ELLIS-MONAGHAN J. A. and MOFFATT I., *Adv. App. Math.*, **47** (2011) 772.
- [40] FORTUNATO S. and BARTHÉLEMY M., *Proc. Natl. Aca. Sci. U.S.A.*, **104** (2007) 36.
- [41] KUMPULA J. M., SARAMÄKI J., KASKI K. and KERTÉSZ J., *Euro. Phys. J. B*, **56** (2007) 41.
- [42] LANCICHINETTI A. and FORTUNATO S., *Phys. Rev. E*, **84** (2011) 066122.
- [43] HU D., RONHOVDE P. and NUSSINOV Z., *e-print arXiv:1204.4167*, (2012).
- [44] BATISTA C. D. and NUSSINOV Z., *Phys. Rev. B*, **72** (2005) 045137.
- [45] NUSSINOV Z., ORTIZ G. and COBANERA E., *e-print arXiv:1110.2179*, (2011).
- [46] JACKSON B. and SOKAL A. D., *J. Combinatorial Theory, Series B*, **99** (2009) 869.
- [47] MITTAG L. and STEPHEN M. J., *J. Phys. A: Math. Nucl. Gen.*, **7** (1974) L109.
- [48] BISKUP M., CHAYES L. and CRAWFORD N., *J. Stat. Phys.*, **122** (2006) 1139.
- [49] PEARCE P. A. and GRIFFITHS R. B., *J. Phys. A: Math. Gen.*, **13** (1980) 2143.
- [50] XIANG J. and HU K., *Physica A*, **391** (2012) 4995.